

The brain as a blueprint: a survey of brain-inspired approaches to learning in artificial intelligence

Guillaume Eter¹

¹Groningen Institute for Evolutionary Life Sciences, University of Groningen, Nijenborgh 7, 97473 Groningen, the Netherlands

¹Correspondance: g.etter@rug.nl

November 7, 2025

Abstract

Inspired by key neuroscience principles, deep learning has driven exponential breakthroughs in developing functional models of perception and other cognitive processes. A key to this success has been the implementation of crucial features found in biological neural networks: neurons as units of information transfer, non-linear activation functions that enable general function approximation, and complex architectures vital for attentional processes. However, standard deep learning models rely on biologically implausible error propagation algorithms and struggle to accumulate knowledge incrementally. While, the precise learning rule governing synaptic plasticity in biological systems remains unknown, recent discoveries in neuroscience could fuel further progress in AI. Here I examine successful implementations of brain-inspired principles in deep learning, current limitations, and promising avenues inspired by recent advances in neuroscience, including error computation, propagation, and integration via synaptic updates in biological neural networks.

1 Introduction

Biological systems adapt to their environments through evolution. This form of adaptation occurs randomly, through genetic permutations, and can be found in the simplest unicellular animals. Lifelong learning, on the other hand, refers to the act of improving behaviors throughout one's life. This process is thought to be directly related to brain activity and involves updating synapses, the points of contact between neurons. In humans, learning can lead to fast improvements in motor (e.g. learning to play tennis) or cognitive (e.g. mastering the game of chess, learning a new language, ...) abilities.

In recent years, artificial intelligence (AI) has come a long way in replicating learning for these cognitive tasks, including vision [1], [2], speech [3], language [4], and problem solving [5]. This success has been fueled by deep learning frameworks [6], which implements key elements from biological neural networks: (1) neurons as units of information transfer [7], [8], (2) nonlinear "all-or-none" [9], [10] activation functions [11], (3) deep architectures that can act as universal function approximators, and (4) synaptic weights as learnable parameters (Fig. 1).

It is remarkable that implementing these principles from biological neural networks is sufficient to design models capable of emulating advanced aspects of human cognition. After all, these models omit much of the complexity observed in the central nervous system, including intricate

architectures and various functional cellular types, from inhibitory interneurons to glial cells.

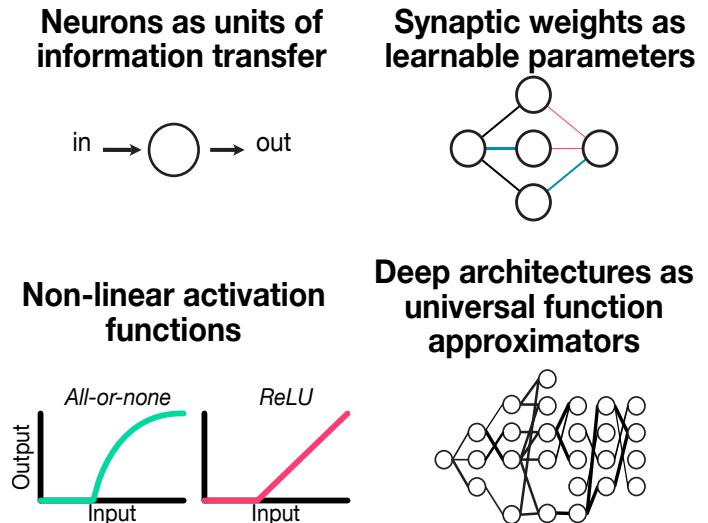


Figure 1: **Brain-inspired principles of deep learning.** Deep learning implements four key principles from neuroscience: neurons as units of computation, synapses as learnable parameters, non-linear activation functions, and deep architectures that act as universal function approximators.

Beyond the aforementioned basic principles of deep

learning, progress in machine learning is still consistently fueled by breakthroughs in neuroscience. For example, convolutional neural networks were inspired by the architecture of the mammalian visual system [12], specifically by implementing hierarchical feature extraction from input patterns [1], [13], [14], [15], leading to visual models that could scale to large and complex images (**Fig. 2**). More recently, vision transformers implement attention mechanisms that are essential in human perception [2].

Hierarchy of feature-specific convolutions

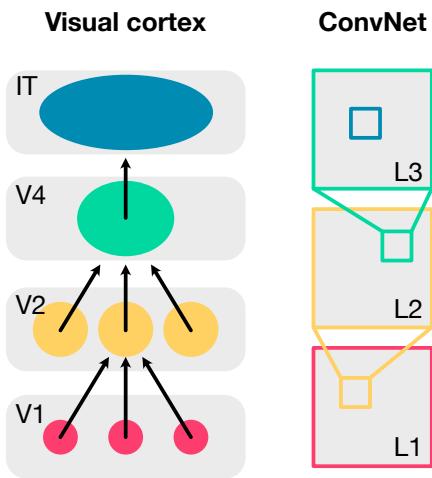


Figure 2: Convolutional neural networks take inspiration from the visual system. The visual system is composed of a hierarchy of feature-selective neural networks that can be composed to learn complex visual representations. Convolutional neural networks take a direct inspiration from this principle by using convolutions in place of one-to-one connectivity patterns found in multilayer perceptrons.

While there is also a long history of unsuccessful attempts at implementing neuroscience principles into artificial models, historically most breakthroughs in AI are somewhat inspired by aspects of the brain. In parallel, this also suggests that failures to implement aspects of the brain into AI models often reflects limits in our understanding of biological computations. This prompts a need to regularly evaluate current advances in neuroscience, and translate them into artificial models to not only advance machine learning but also test theories of brain function.

Here, I review recent advances in neuroscience, including *in vivo* plasticity mechanisms, interneuron microcircuits, oscillations, and the potential role of offline replay in continual learning. Throughout this exploration, I review attempts at implementing these advances in machine learning and what can be learned from the success and failures of these experimentations. Finally, I briefly cover emerging frameworks including neuromorphic hardware, which has the potential to fuel future implementations of biologically plausible learning algorithms directly into physical hardware.

2 Towards advanced bioinspired neural networks

Learning can be broken down into distinct problems including: (1) error computation and propagation, (2) credit assignment, (3) synaptic updates. This breakdown is mostly useful for reviewing advances in these areas, nevertheless it is noteworthy that these problems are often entangled in practice. Later, I will also cover the problem of continual learning where the objective is to train neural networks on new tasks continuously without forgetting previous tasks.

2.1 Computing errors locally with inhibitory interneurons

Crucially, biological and artificial systems require *some* feedback to know whether behavioral outputs are appropriate for a given task. This is referred to as the error signal. If an error signal is zero, it is assumed that the system (model or agent) is performing optimally and no synaptic updates need to be performed. When nonzero, the error signal provides a feedback for how much worse output behaviors are from some target (or objective). In machine learning, this objective is often explicit in supervised training contexts, while in biological systems, the objective may only be survival, and error signals could take the form of physiological signaling and/or more specialized signals. These error signals can be global or local, and are crucial to providing credit assignment during learning. In the brain, it was proposed that dopamine release encodes temporal difference errors, a form of global error signal [16]. Such global error signals are likely not the main drivers of credit assignment [17], but could be used in conjunction to feedback signals and local inhibitory drive (see below) to gate and set the sign of synaptic updates [18]. Beyond dopamine which has been fairly well described in the brain and will not be covered here, local error computations (at the level of a single hierarchical level, or microcircuit) remain elusive. Here, we can examine the main classes of interneurons in the brain, how they interact within a microcircuit, and how they could contribute to computing local error signals.

A key distinction between artificial and biological neural networks is that in the former, all units are typically comparable in function, whereas in the latter, some neurons release excitatory neurotransmitters while others release inhibitory neurotransmitters. This dichotomy between excitatory and inhibitory neurons is referred to as 'Dale's principle' [19], which ANNs typically do not respect.

The nomenclature of interneurons is complex, typically defined both by function, morphology, and genetic expression, and encompasses hundreds of subtypes [20], [21]. In spite of this broad diversity, main functional classes have emerged in the past few years.

2.1.1 Parvalbumin interneurons

First, parvalbumin-expressing (PV) interneurons represent the largest class of cortical inhibitory cells. They are typically fast-spiking and primarily target the perisomatic region (cell body and proximal dendrites) of principal neurons [22]. This strategic targeting allows PV cells to exert powerful control over the output spiking activity of principal, excitatory neurons. Functionally, PV interneurons are critical for establishing network stability through strong feedback and feedforward inhibition, preventing runaway excitation [23]. They also play a key role in synchronizing neural populations in theta (around 8 Hz) [24] and gamma (30 - 120 Hz) [25], frequency bands, which is thought to be instrumental for temporal coding, feature binding, and routing information across brain regions [26], [27]. By enforcing precise spike timing and controlling overall network gain, PV interneurons create a stable yet dynamic environment conducive to reliable neural computation. Importantly, PV neurons also directly gate learning, and their activity decreases upon reaching goals in goal-directed tasks [28].

2.1.2 Somatostatin interneurons

In contrast to the perisomatic targeting of PV cells, somatostatin-expressing (SOM) interneurons primarily innervate the dendrites of principal neurons, particularly the apical tufts where many excitatory inputs arrive [29], [30], [31]. SOM neurons typically exhibit adapting firing patterns and respond well to persistent activity. Their dendritic inhibition allows them to selectively filter or gate synaptic inputs, influencing dendritic integration and the conditions required for triggering dendritic spikes, which are themselves powerful computational events [32]. By controlling dendritic excitability, SOM interneurons can modulate synaptic plasticity occurring at dendritic synapses. They essentially act as a gating mechanism on dendritic activity and associated plasticity [33]. Their potential role in learning is even more striking when considered together with VIP interneurons.

2.1.3 VIP interneurons

Vasoactive Intestinal Peptide-expressing (VIP) interneurons constitute a third major class, unique in that they predominantly target other interneurons, most notably SOM cells [34]. This creates a feedforward disinhibitory circuit: activation of VIP neurons inhibits SOM neurons, which in turn reduces the inhibition onto the dendrites of principal neurons. This disinhibition effectively opens a gate for dendritic activity and plasticity (which I will cover in more details below). VIP neurons are heavily modulated by top-down signals, including reinforcement signals and attentional cues, often conveyed by neuromodulators like acetylcholine and norepinephrine [35], [36]. Therefore, the VIP-SOM-Pyramidal cell circuit provides a mechanism for context-dependent gating of learning (Fig. 3). When behaviorally relevant signals activate VIP neurons, SOM inhibition of the apical shaft is temporarily lifted, allowing

synapses to be modified based on local dendritic computations, which could represent prediction errors or relevant associations [37].

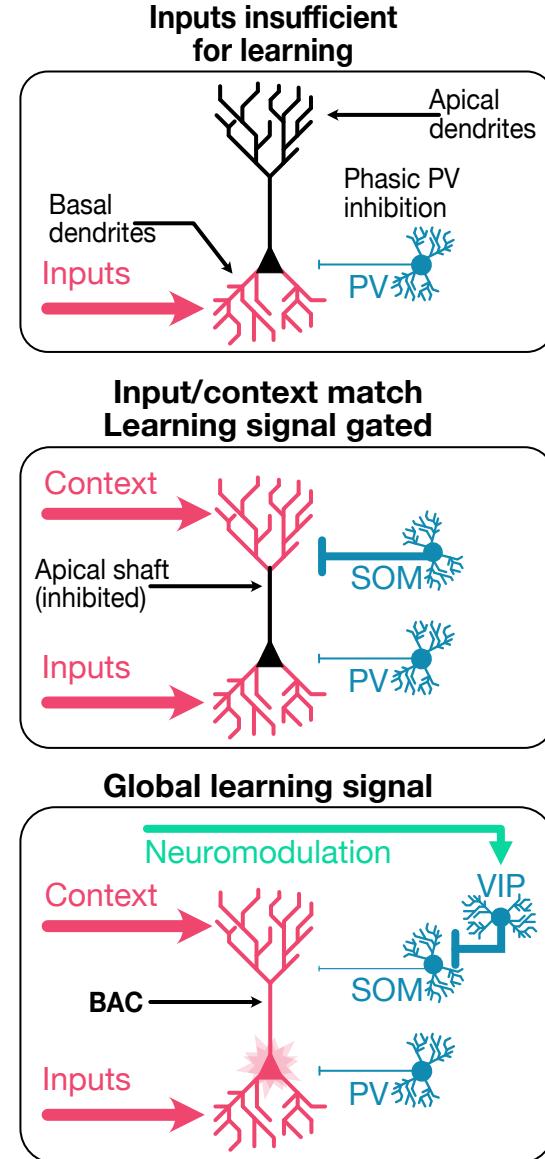


Figure 3: Inhibitory microcircuit. Top, parvalbumin (PV) neurons contribute to maintaining neurons in an equilibrium state, whereby weak activation of basal dendrites by inputs is insufficient to elicit output action potentials. Middle, when apical dendrites are activated by contextual inputs, somatostatin (SOM) neurons prevent backpropagation of calcium (BAC) signals by inhibiting the apical shaft. Bottom, when global learning signals are provided through neuromodulation, vasoactive intestinal peptide (VIP) provide feed-forward disinhibition of the apical shaft and allow BAC signals to propagate.

2.1.4 Implementing inhibitory interneurons in ANNs

It is noteworthy that the power of these microcircuits may lie in the combined integration of global (neuromodulatory) and local error signals. This also suggests that these two

types of error signals may not be disentangled in biological neural networks, while they rarely co-exist in current AI models. While computational models of these inhibitory interactions have been proposed [38], integrating these concepts in deep ANNs remains an active area of research. Remarkably, ANNs that respect Dale's principle can be trained with the same efficiency as regular non-Dalean ANNs as long as inhibition centres and standardizes the excitatory signals [39]. This principle can also be extended to recurrent neural networks, further supporting the potential for alignment with biological neural networks [40].

2.2 Biologically plausible credit assignment

Assuming that a useful error signal is available, the next problem is to identify which parameters need to be updated to improve output behaviors. Credit assignment is a key principle in learning and can be understood at different levels of computations. The simplest analogy is that of a game of chess, where credit assignment refers to assigning which move(s) contributed most to winning or losing the game. This concept is also applied directly to units of computations, neurons, at inference time: i.e. which neurons contributed most to successful decisions and adaptive behaviors. As learning itself involves changes at the synaptic levels though, credit assignment is most useful when considered as identifying which parameters most contribute improving performance. Credit assignment is one of the most active yet challenging areas of research in machine learning and neuroscience. As mentioned above, an interaction between feedback and neuromodulatory signals could effectively provide credit assignment in the brain [18]. Currently however, standard deep learning systems still rely of backpropagating error gradients throughout the entire network, which is famously biologically implausible [41], [42]. While effective in many applications, there are several reasons for which alternatives to backpropagation, as well as establishing clearly what credit assignment mechanisms are used in the brain, could be valuable.

2.2.1 Backpropagation

Backpropagation (BP) is efficient at computing gradients and updating weights in multi-layered networks and has many advantages over other learning algorithms in terms of scalability and generalization [41], [42]. However, BP is famously biologically implausible [17], [43], [44]. This implausibility has been known since its inception and extensively described in the literature, and here I will only briefly summarize the main points. Firstly, BP requires the computation of exact gradients of the loss function with respect to the weights, which is not how learning is believed to occur in biological systems. Secondly, BP relies on the assumption of symmetric weights, where the weights used for forward propagation are the same as those used for backward propagation. One major issue is the "weight transport problem," where BP requires the error signals to be propagated backward through the network using the

exact transpose of the weights used in the forward pass [17], [45], [46]. This requirement of precise symmetry between forward and backward connections is not supported by biological evidence, as synapses in the brain are generally unidirectional. Thirdly, BP requires that error signals at each layer is computed based on the output of the entire network, which is not how learning occurs in biological systems [17], [43], [44]. Finally, BP explicitly separates training and inference in ways that is unlikely to occur in the brain. I will cover this final point separately when looking at continual learning. While BP still remains the state-of-the-art approach to train ANNs, several alternatives have shown promising results.

2.2.2 Feedback alignment

One core idea to implement credit assignment without BP is to provide feedback with a dedicated network. Early attempts include Feedback Alignment (FA) [47] is one such approach that offers a surprisingly simple yet effective way to circumvent the need for weight symmetry. The core idea of FA is to use a fixed random matrix for the feedback pathway instead of the transpose of the feedforward weight matrix. The feedforward weights then learn to adjust themselves such that the error signal propagated through the random feedback matrix provides a useful direction for learning. Despite the seemingly arbitrary nature of the feedback weights, FA has shown to be surprisingly effective on various tasks, suggesting that precise weight symmetry might not be essential for learning in deep networks.

2.2.3 Sign symmetry

Another another approach is to relax the weight symmetry requirement by using the sign of the feedforward weights for the feedback pathway, rather than the exact values [46]. Here the magnitude of the feedback weights does not need to match the forward weights; only the sign is shared. This approach is biologically more plausible than strict weight symmetry, as it only requires the direction of influence (excitatory or inhibitory) to be consistent between the forward and backward paths. Interestingly, networks trained with sign symmetry have also demonstrated promising performance on challenging datasets like ImageNet, approaching the accuracy of backpropagation in some cases. The success of feedback alignment and sign symmetry highlights that relaxing the strict weight symmetry constraint of backpropagation while still providing a meaningful error signal can lead to effective learning in deep neural networks.

2.2.4 Target propagation

Instead of propagating error gradients backward, Target Propagation (TP) suggests propagating target values for the neural activations [48], [49], [50], [51], [52], [53]. In this framework, each layer aims to achieve a specific target value, and these targets are determined by propagat-

ing information backward from the output layer. A common approach in TP involves using layer-wise autoencoders, where each layer learns to reconstruct its input, and the targets for a layer are generated based on the reconstructed activity of the subsequent layer. The parameters of the feedforward network are then updated locally to move the actual activations closer to these targets. While the concept of target propagation offers a biologically appealing alternative, the initial implementations faced challenges, particularly due to the imperfection of the learned inverse mappings or autoencoders [50], [53], [54].

2.2.5 Difference target propagation

To address these issues and improve the stability and performance of target propagation, Difference Target Propagation (DTP) was introduced [50]. DTP incorporates a linear correction mechanism to the target propagation process. The target for a layer in DTP is not solely based on the output of the feedback network but also takes into account the current activation of that layer and the target of the subsequent layer. This difference correction helps to stabilize the training process and allows DTP to achieve performance levels that are more comparable to backpropagation. The introduction of DTP marked a significant advancement in the field, demonstrating a more robust and effective approach to biologically plausible learning through target-based mechanisms.

2.2.6 Direct difference target propagation

Another direction in the evolution of target propagation has led to the exploration of Direct Difference Target Propagation (DDTP; [54]). In contrast to the layer-by-layer feedback in standard DTP, DDTP investigates the use of direct feedback connections from the output layer to each of the hidden layers. This approach aims to streamline the backward propagation of target information and potentially improve the efficiency of the learning process. Furthermore, advancements in the feedback mechanisms used within target propagation have resulted in the development of specialized loss functions, such as the Local-Difference Reconstruction Loss (L-DRL; [55]). L-DRL is a feedback loss function designed to improve the performance of DTP by facilitating a better alignment between the feedback weights and the effective error signals needed for learning in the feedforward pathway. DTP, when combined with L-DRL, has achieved remarkable results on image classifi-

cation tasks, with performance levels that are increasingly comparable to those of backpropagation on the same architectures. These findings underscore the importance of a feedback mechanisms and loss functions in target propagation to effectively guide learning in a biologically plausible manner.

2.2.7 Fixed-weight difference target propagation

Further refinements of Difference Target Propagation have led to the development of Fixed-Weight Difference Target Propagation (FW-DTP) [56]. A key characteristic of FW-DTP is that it keeps the feedback weights constant throughout the training process. This simplification offers several advantages over standard DTP, where the feedback weights are typically learned alongside the feedforward weights. By fixing the feedback weights, FW-DTP significantly reduces the computational cost associated with training, as it eliminates the need to update the feedback pathway. Despite this simplification, FW-DTP has been shown to still effectively deliver informative target values to the hidden layers of the network, enabling learning to occur. FW-DTP can achieve improved stability during training and often exhibits higher test performance compared to standard DTP on various image classification datasets. This suggests that the core concept of target propagation, particularly with the difference correction mechanism, can function effectively even without the need for a learned feedback pathway. The success of FW-DTP indicates that optimizing the objectives of layer-wise target reconstruction, which often involves training feedback weights, might not be strictly necessary for the concept of target propagation to work properly.

These algorithms based on target propagation highlight the potential role of feedback connections. It is noteworthy that in the neocortex and throughout most brain regions, neural networks form not only feed-forward, but also feedback (top-down) projections. Multiple roles have been attributed to these top-down projections, including attention [57], perception [58], representing predictions [59], [60] Beyond credit assignment during learning, bidirectional architectures (that take inspiration from the neocortex and include both bottom-up and top-down pathways) have significant benefits in representation learning. Specifically, top-down signals can be leveraged to resolve ambiguity given that contextual information is available [61]. Similarly, context-invariant representations can be learned by feedback rather than feedforward processing [62].

Algorithm	Achievement	Main Limitation
Feedback Alignment (FA)	A biologically plausible alternative to backpropagation that uses fixed, random backward weights to align with the forward weights.	Performance can be slightly lower than backpropagation on complex tasks; convergence can be sensitive to network architecture and initialization.
Sign Symmetry	A biologically plausible alternative to backpropagation, leveraging symmetric error signs rather than weight for credit assignment.	Requires careful initialization and training; convergence can be sensitive to hyperparameters.
Target Propagation (TP)	Addresses the non-local credit assignment problem in deep networks by propagating "targets" from the output layer backward.	Can be complex to implement; the "inverse" function required for target propagation might be difficult to learn or define for arbitrary layers.
Difference Target Propagation (DTP)	An extension of TP that focuses on minimizing the difference between the actual and target activations, simplifying the target generation.	Still relies on the concept of inverse operations, which can be challenging to learn accurately, especially in highly non-linear networks.
Direct Difference Target Propagation (DDTP)	A variant of DTP that aims to avoid explicitly learning inverse functions by directly training a "backward" network.	Can still require careful tuning of the backward network; the backward pathway needs to be learned effectively.
Fixed-Weight Target Propagation (FWTP)	Explores target propagation with fixed backward weights (or a fixed backward pathway), simplifying the learning process.	The fixed backward weights might limit the flexibility and representational power compared to learned backward pathways; performance can be sensitive to the choice of fixed weights.

Table 1: **Alternatives to backpropagation.** Comparison of algorithms that do not rely on error gradient backpropagation and present biological plausibility.

2.3 Behavioral timescale plasticity and burst-prop

Assuming that an appropriate error signal can be derived during learning and that credit assignment can be performed appropriately, the next problem is to find an effective approach to update synapses. Note that in effect, credit assignment and synaptic updates are not completely separable. In his seminal work [63], Donald Hebb posited that when two neurons are repeatedly activated in close temporal proximity, the connection between them is strengthened. This principle, often summarized as "neurons that fire together, wire together," suggested a cellular mechanism by which experiences could induce lasting changes in neural circuits, forming the basis of learning and memory.

2.3.1 Long-term potentiation

Experimental evidence supporting Hebbian-like plasticity emerged decades later with the discovery of long-term potentiation (LTP) [64]. In these experiments, a high-frequency train of stimuli applied to a neural pathway could

lead to a long-lasting increase in the strength of synaptic transmission for that pathway. While LTP demonstrated that synaptic strength could be modified based on activity, it was a relatively coarse measure of correlated firing.

2.3.2 Spike-time-dependent plasticity

The next significant development was the concept of spike-time-dependent plasticity (STDP). STDP refined the Hebbian principle by demonstrating that the relative timing of pre- and postsynaptic action potentials was critical in determining the direction and magnitude of synaptic plasticity. Specifically, if a presynaptic spike consistently precedes a postsynaptic spike within a narrow time window, the synapse is potentiated (LTP). Conversely, if the postsynaptic spike precedes the presynaptic spike, the synapse is depressed (LTD) [65], [66]. This timing-dependent rule offered a more nuanced mechanism for how neurons could learn about the temporal relationships between their inputs and their own firing.

One limitation is that the standard STDP model primarily operates on a millisecond timescale, determined by the precise timing of individual spikes. However, many forms

of behavioral learning and memory formation occur over much longer timescales, ranging from seconds to minutes. It has been questioned whether millisecond-precision STDP alone can fully account for plasticity engaged during complex behaviors [67]. Additionally, the strict requirement for precisely timed pre- and postsynaptic spikes might be less robust in the noisy and asynchronous environment of a living brain. Finally, these principles have largely been established in the *in vitro* (brain slices) context and may not translate *in vivo* as (1) they require a high number of stimulus-response pairings, which are typically not possible in fast-learning contexts, and (2) calcium concentration, and thus neurotransmitter release probabilities typically differ between *in vivo* and *in vitro* settings [68], [69].

2.3.3 Behavioral timescale plasticity

These limitations have paved the way for the investigation of plasticity mechanisms operating on longer timescales, leading to the emergence of the concept of behavioral timescale plasticity (BTSP) [70], [71]. BTSP proposes that synaptic plasticity can be driven by events occurring over seconds, a timescale more aligned with natural behaviors and learning processes. While the precise mechanisms of BTSP are still being elucidated, it is thought to involve the integration of synaptic inputs with longer-lasting postsynaptic depolarizations, such as plateau potentials, which can gate plasticity over extended periods [70], [71]. BTSP is particularly interesting for several reasons. Firstly, its operating timescale aligns well with the temporal dynamics of behavioral learning, providing a more plausible link between neural plasticity and the formation of memories in a naturalistic setting. Secondly, BTSP in hippocampal neurons has been shown to be crucial for the formation of stable spatial representations (place fields) during navigation, even with single learning trials [70], [72]. This suggests a mechanism for rapid, one-shot learning that is difficult to explain with standard millisecond-based STDP. Finally, recent research indicates that BTSP can also contribute to the formation of non-spatial representations, suggesting a more general role in hippocampal function beyond spatial coding [73]. BTSP represents a significant step towards understanding how neural circuits can integrate information over behaviorally relevant timescales to support learning and memory.

2.3.4 Role of backpropagating associated calcium spikes in plasticity

It is also important to briefly mention the cellular and physiological principles behind these plasticity events, which will have implications down the line when extrapolating learning algorithms from cellular interactions. Backpropagating action potentials (BAPs) are electrochemical signals that travel from a neuron's soma back into its dendritic tree [74]. This backpropagation actively depolarizes dendritic branches, though the extent can vary between basal and apical dendrites, with more complex dynamics in the distal

apical tuft [75]. Calcium ions (Ca^{2+}) are critical for synaptic plasticity and other functions [76]. A key mechanism linking electrical activity to plasticity is the coincidence detection of synaptic input (EPSPs) and postsynaptic BAPs. This co-occurrence leads to sufficient depolarization to relieve the magnesium block of NMDA receptors, allowing Ca^{2+} influx [32], [66]. The precise timing of the EPSP relative to the BAP determines the magnitude and duration of this calcium signal, underpinning plasticity mechanisms such as STDP [65].

In addition to the calcium entry directly associated with BAP depolarization, the conjunction of strong synaptic input and BAPs, particularly in apical dendrites, can trigger regenerative backpropagating associated calcium (BAC) spikes [77]. These BAC spikes involve substantial calcium influx through both NMDA receptors and voltage-gated calcium channels, generating a larger and more sustained calcium transient within the dendrite [78]. The amplitude and temporal profile of this intracellular calcium signal act as a crucial switch, determining the specific form of plasticity induced [79].

Ultimately, calcium signals generated by BAPs and BAC spikes translate into changes in synaptic strength by activating different downstream signaling pathways. High levels of calcium tend to activate protein kinases such as CaMKII, which promote the insertion of AMPA receptors into the postsynaptic membrane, leading to LTP [80]. These activity-dependent modifications of synaptic efficacy are fundamental cellular processes that enable neural circuits to adapt and are essential for learning and memory. In the next section, I will briefly review how these known physiological principles can be translated into learning rules for artificial systems.

2.3.5 Silence, single, and bursting activity: ternary code of biological neural networks

Neurons in the brain do not always fire individual spikes; they can also fire in rapid sequences of action potentials known as bursts [81]. Burst firing can induce more significant changes in the strength of synaptic connections compared to isolated single spikes. This phenomenon, known as burst-dependent plasticity, suggests that bursts play a crucial role in learning and information processing in the brain. In effect, this suggests that rather than a binary code, neurons could leverage a ternary code of silence, bursts, and single action potentials [82]. The biological significance of neuronal bursts as a mechanism for learning and communication has inspired the development of burst-propagation learning algorithms in ANNs, aiming to leverage these principles for more biologically plausible and potentially more powerful learning methods.

Building upon the biological evidence of burst-dependent plasticity, bursting activity has been proposed as a candidate mechanism to provide temporal credit assignment [83]. One formalism of this idea is Burstprop, which introduces regularizing feedback connections for each neuron and the use of two types of activity patterns

to communicate signed error information [84]. This allows Burstprop to be applied to tasks like image classification on the MNIST dataset using networks with hundreds of neurons, a significant step towards handling more realistic problems. Notably, Burstprop utilizes feedback alignment methods for transporting error signals backward through the network, which is a biologically plausible way to address the weight transport problem. Experimental results have shown that Burstprop can achieve low test classification errors on MNIST, with performance levels that are comparable to those obtained using backpropagation through time on the same network architectures. This demonstrates the potential of burst-dependent learning, as implemented in Burstprop, to scale to more complex learning tasks and serve as a viable method for learning with neuromorphic hardware (which we will briefly cover later).

An intriguing aspect of Burstprop is that it can approximate error backpropagation, effectively performing a similar form of credit assignment but through a more biologically plausible mechanism. In this context, neuron bursts can be viewed as acting as teaching signals, conveying information about the error in the network's output, similar to how error signals are propagated in backpropagation.

2.4 The continual learning challenge

Training in deep learning is usually done in batches and separated entirely from inference, where synaptic weights remained fixed. While for many practical applications, this separation is mostly desirable, it poses problems in the context of continual learning, where learning systems are expected to accumulate knowledge over various tasks and contexts without forgetting previous knowledge. At the end of training, deep learning models should not only perform well on input close to the training set, but also on never-seen before input, provided that it is part of the underlying distribution learned during training. Out-of-distribution inputs are usually not well processed in these conditions, and require additional training to extend the learned distribution to new scenarios. This currently still poses a significant challenge for current neural networks, as re-training a model on new tasks will usually lead to the loss of previous knowledge - a phenomenon commonly referred to as 'catastrophic forgetting' [85], [86], [87]. This section is not meant to cover the continual learning problem extensively, but rather mention relevant brain-inspired approaches. For a more extensive account, see Wang et al. [88].

2.4.1 Continual learning with structural changes

One set of mitigation approaches focuses on promoting sparse representations, where only a small fraction of neurons are active for any given input, including dropout [89] or the k-Winner-Take-All (k-WTA) activation function, which encourage such sparsity [90]. Another approach is to embed some form synaptic stability directly into the model. In the brain, synapses are situated on dendritic spines.

These spines undergo distinct morphological stages during plasticity: thin filopodial spines are new and unstable, whereas mushroom-shaped spines are stable and less likely to decay [91]. In deep learning, approaches like Synaptic Intelligence [92] or elastic weight consolidation [93] aim to protect important synapses from being overwritten when learning new tasks. Surprisingly, implementing biologically-inspired synaptic turnover mitigates the progressive loss of plasticity observed in connectionist models learning continually [94].

2.4.2 Continual learning with replay

Another important class of solutions for continual learning takes direct inspiration from replay. While animals are believed to learn continuously, this does not mean that training and inference completely overlaps, as mammals go through sleep-wake cycles that are associated with distinct brain states, including REM and slow-wave sleep. In animals, including humans, neurons tend to replay awake patterns during sleep [95], [96], [97]. While replay has traditionally been associated with memory consolidation [98], [99], [100], [101], its exact physiological role remains unknown. Additionally, replay occurs during wakefulness, where it could have roles beyond consolidation. One recent hypothesis is that of awake replay improving future goal-oriented behavior by internally simulating past experiences, then using them to adjust estimates of future rewards, effectively biasing neural activities towards prospective rewards [102] (**Fig. 4**). Congruent with this hypothesis has been the application experience replay to encourage ANNs to maintain similar predictions for the replayed samples [90]. While most experience replay approaches either directly memorize previous samples [103], [104], [105], [106], [107], [108], [109], [110], [111], [112] or generate previous samples in input space [113], [114], [115], [116], [117], [118], [119], it is also possible to generatively replay previous samples in latent space [120], which is more biologically plausible.

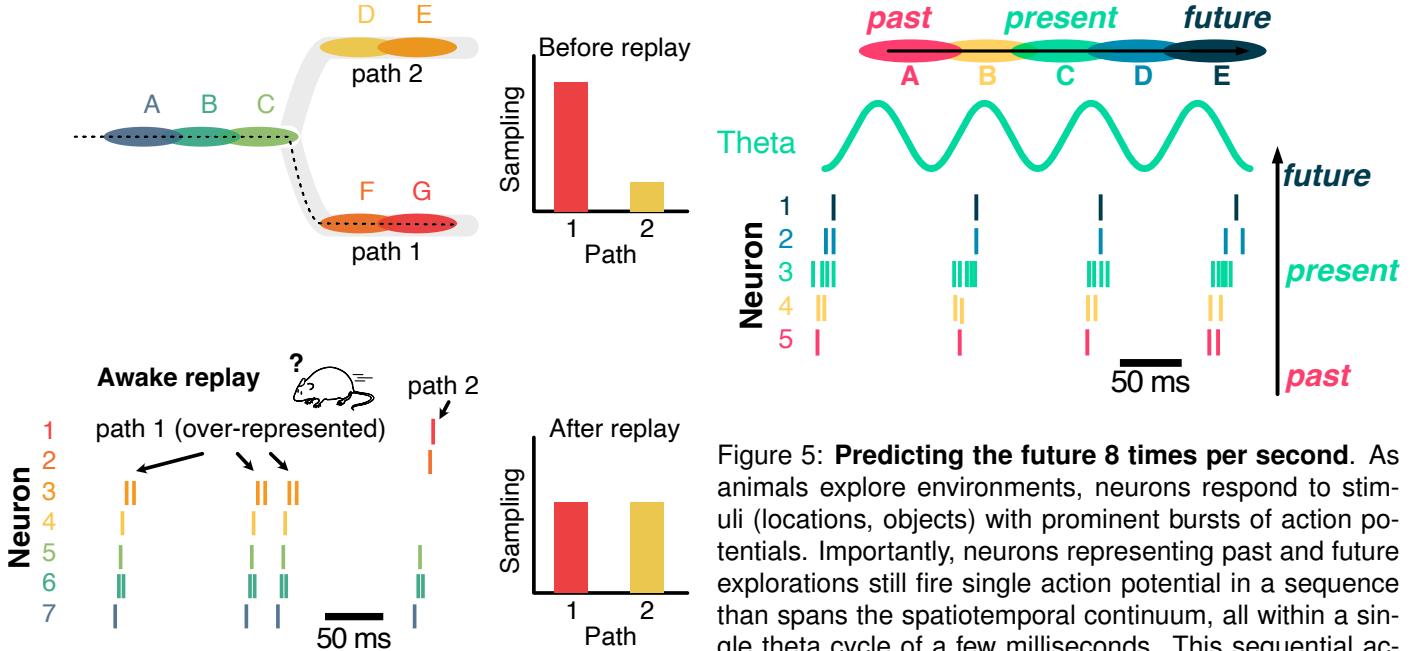


Figure 4: Role of replay in continual learning. Neuronal activities associated with awake exploration are replayed during periods of inactivity (quiet wakefulness). While replay has often been associated with memory consolidation (during sleep), an alternative hypothesis is that representing under-sampled experiences in an offline state could correct for biases in the 'training data' of sampled explorations [102], [121].

2.4.3 Continual learning in real-time

This section will explore more speculative roles of known physiological mechanisms in continual learning. *In vivo*, the mammalian brain expresses a wide range of oscillations that reflect the coordinated activity of excitatory and inhibitory neurons [25], [27]. Importantly, the excitability of individual neurons is modulated by these broader oscillation signals. Theta oscillations (6 - 12 Hz) in particular may provide temporal windows during which plasticity can occur [121]. During awake exploration, neurons representing locations (or objects) tend to do so in sequences that span the past, present, and potential futures. Neurons representing present exploration tend to burst, whereas neurons representing past and future experiences fire single action potentials (**Fig. 5**). These so called theta-sequences could introduce robustness against catastrophic forgetting, as both immediate memory and future planning can co-exist quasi-simultaneously during active cognition [121].

Figure 5: Predicting the future 8 times per second. As animals explore environments, neurons respond to stimuli (locations, objects) with prominent bursts of action potentials. Importantly, neurons representing past and future explorations still fire single action potential in a sequence than spans the spatiotemporal continuum, all within a single theta cycle of a few milliseconds. This sequential activity reflects the maintenance of past-present-future representations quasi-simultaneously, and could be relevant for learning new information without forgetting immediate past experiences [121].

When considered together with the aforementioned burst-based learning rules and inhibitory interneuron microcircuits, local field and dendritic oscillations could provide a direct mean for testing future predictions, learning relevant information, while maintaining previous immediate experiences, all in real-time.

The main idea behind this tenet is that distinct basal and apical dendritic compartments oscillate within the theta frequency band, but at distinct oscillatory phases. Calcium can accumulate specifically in apical dendrites [122], and concomitant basal and apical activations can drive the generation of plateau potentials reliably [123], [124]. This suggests that in addition to an upstream teaching signal that results from the interaction between interneurons, neuromodulators, and plateau potentials, there are specific temporal windows during which new information can be encoded, as determined by the underlying dendritic oscillations. Given that theta sequences contain information about immediate past experiences and future goals, and input representations co-exist in dendrites [125], [126], this could provide a transient mechanism to support continual learning in real time (**Fig. 6**).

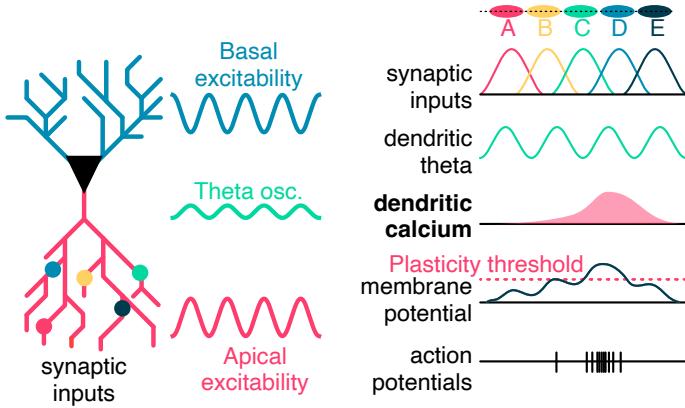


Figure 6: Real-time calcium-based plasticity. In addition to the interneuron microcircuit discussed earlier, excitatory neurons fire with respect to underlying oscillations. Remarkably, these oscillations are differentially expressed in dendrites, and can contribute to the exact timing of output action potentials. In this model, I propose that past, current experiences, and future decisions are all encoded in distinct synaptic inputs and expressed sequentially as sub-threshold activity in dendrites. Only when future predictions are being tested at present time, can dendritic calcium increase to levels that enable bursts of action potentials, synaptic plasticity, and learning [121].

Neuromorphic computing

The premise of neuromorphic computing is that the brain's architecture and processing principles represent a highly efficient and effective blueprint for certain types of computation, particularly those involving pattern recognition, learning, and adaptation in complex, noisy environments [127]. Current AI models, predominantly based on deep learning running on conventional von Neumann architectures (with separate processing and memory units), has achieved remarkable success but faces significant hurdles. These include immense energy consumption, the need for vast datasets for training, and limitations in continuous learning and robustness to novel situations. Neuromorphics proposes a paradigm shift, aiming to build hardware that directly mimics neural structures and dynamics, potentially overcoming these limitations by leveraging the brain's inherent parallelism, event-driven processing, and co-location of memory and computation [128].

Memristors, or resistive memory devices, have garnered significant attention as potential building blocks for neuromorphic synapses due to their ability to change resistance based on the history of voltage applied or current passed through them, and retain that state [129], [130]. Their key advantages include nanoscale size (allowing for extremely dense memory arrays), analog (multi-level) resistance states (ideal for representing synaptic weights), non-volatility (retaining memory without power), and the potential for integration into crossbar arrays that naturally perform vector-matrix multiplications crucial for neural networks, thus enabling efficient in-memory computing

and mitigating the von Neumann bottleneck [131]. Basic STDP-like behavior, pattern classification, and other useful synaptic functions have been successfully implemented using memristor arrays [132], [133].

Despite the promise, significant challenges remain. Memristive devices often suffer from considerable variability (device-to-device and cycle-to-cycle fluctuations in behavior), limited endurance (number of times they can be reliably switched), non-ideal switching characteristics (non-linearity, asymmetry between potentiation and depression), and difficulties in achieving precise, incremental weight updates needed for many learning algorithms [128], [134]. Integrating these nanoscale devices reliably with standard CMOS circuits for control and readout also poses manufacturing challenges. Overcoming these hurdles is critical for realizing large-scale, reliable memristor-based neuromorphic systems.

One of the most compelling arguments for neuromorphic computing is its potential for energy efficiency. Training state-of-the-art deep learning models, especially massive language models, requires enormous computational resources, consuming megawatts of power for weeks or months, resulting in substantial carbon footprints [135], [136]). The operational (inference) costs are also significant, particularly as AI becomes more pervasive. This energy demand poses a serious challenge to sustainable AI development and limits the deployment of complex AI on power-constrained devices (e.g. mobile phones, wearables, remote sensors). Neuromorphic systems, through their event-driven operation (computation only where needed), inherent parallelism, sparse activity, and co-location of memory and processing (reducing costly data movement), promise energy savings potentially orders of magnitude greater than conventional hardware for equivalent tasks [137]. Achieving this efficiency could reduce the environmental impact of data centers and enable entirely new applications requiring continuous, low-power intelligent processing.

Conclusions

Advances in deep learning have been in large part enabled by breakthroughs in neuroscience. Conversely, deep learning frameworks have provided neuroscientists with computational frameworks to test fundamental questions about how the brain might operate. Beyond well established neuroscience principles in deep learning, such as nonlinearities and architectures, many recent neuroscience discoveries must be brought to the attention of the machine learning community. In particular, recent discoveries related to behavioral timestamp plasticity, inhibitory interneuron microcircuits, and real-time *in vivo* computations with oscillations could provide testing grounds for the next generation of AI systems.

Critically, von Neumann hardware will eventually be limited in its ability to implement biologically plausible artificial systems, in particular those that rely on spike-based learn-

ing rules. Ultimately, progress for these biologically plausible algorithms is limited by the very knowledge frontier in neuroscience, and neurophysiological experimentation still holds significant value for AI practitioners. While still an emerging field, neuromorphics present a valuable potential for running biological computations while circumventing the limitations of von Neumann hardware all the while reducing energy demand, thus tackling some of the biggest challenges ahead of machine learning.

References

- [1] Y. LeCun and Y. Bengio, “Convolutional networks for images, speech, and time series,” *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.
- [2] A. Dosovitskiy et al., *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*, 2021. DOI: [10.48550/arXiv.2010.11929](https://doi.org/10.48550/arXiv.2010.11929) [cs]. arXiv: [2010.11929](https://arxiv.org/abs/2010.11929) [cs].
- [3] A. van den Oord et al., *WaveNet: A Generative Model for Raw Audio*, 2016. DOI: [10.48550/arXiv.1609.03499](https://doi.org/10.48550/arXiv.1609.03499) arXiv: [1609.03499](https://arxiv.org/abs/1609.03499) [cs].
- [4] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, “Improving Language Understanding by Generative Pre-Training,”
- [5] D. Silver et al., “Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, no. 7587, pp. 484–489, 2016, ISSN: 1476-4687. DOI: [10.1038/nature16961](https://doi.org/10.1038/nature16961)
- [6] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, ISSN: 14764687. DOI: [10.1038/nature14539](https://doi.org/10.1038/nature14539)
- [7] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943, ISSN: 1522-9602. DOI: [10.1007/BF02478259](https://doi.org/10.1007/BF02478259)
- [8] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1958, ISSN: 1939-1471. DOI: [10.1037/h0042519](https://doi.org/10.1037/h0042519)
- [9] K. Lucas, “The “all or none” contraction of the amphibian skeletal muscle fibre,” *The Journal of Physiology*, vol. 38, no. 2-3, pp. 113–133, 1909, ISSN: 0022-3751. DOI: [10.1113/jphysiol.1909.sp001298](https://doi.org/10.1113/jphysiol.1909.sp001298)
- [10] E. D. Adrian and Y. Zotterman, “The impulses produced by sensory nerve-endings,” *The Journal of Physiology*, vol. 61, no. 2, pp. 151–171, 1926, ISSN: 0022-3751. DOI: [10.1113/jphysiol.1926.sp002281](https://doi.org/10.1113/jphysiol.1926.sp002281)
- [11] A. S. Householder, “A theory of steady-state activity in nerve-fiber networks: I. Definitions and preliminary lemmas,” *The bulletin of mathematical biophysics*, vol. 3, no. 2, pp. 63–69, 1941, ISSN: 1522-9602. DOI: [10.1007/BF02478220](https://doi.org/10.1007/BF02478220)
- [12] D. H. Hubel and T. N. Wiesel, “Receptive fields and functional architecture of monkey striate cortex,” *The Journal of Physiology*, vol. 195, no. 1, pp. 215–243, 1968, ISSN: 1469-7793. DOI: [10.1113/jphysiol.1968.sp008455](https://doi.org/10.1113/jphysiol.1968.sp008455)
- [13] K. Fukushima, “Cognitron: A self-organizing multilayered neural network,” *Biological Cybernetics*, vol. 20, no. 3, pp. 121–136, 1975, ISSN: 1432-0770. DOI: [10.1007/BF00342633](https://doi.org/10.1007/BF00342633)
- [14] K. Fukushima, “Neocognitron: A hierarchical neural network capable of visual pattern recognition,” *Neural Networks*, vol. 1, no. 2, pp. 119–130, 1988, ISSN: 08936080. DOI: [10.1016/0893-6080\(88\)90014-7](https://doi.org/10.1016/0893-6080(88)90014-7)
- [15] K. Fukushima, “Neocognitron for handwritten digit recognition,” *Neurocomputing*, vol. 51, pp. 161–180, 2003, ISSN: 09252312. DOI: [10.1016/S0925-2312\(02\)00614-8](https://doi.org/10.1016/S0925-2312(02)00614-8)
- [16] E. J. P. Maes et al., “Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors,” *Nature Neuroscience*, vol. 23, no. 2, pp. 176–178, 2020, ISSN: 1546-1726. DOI: [10.1038/s41593-019-0574-1](https://doi.org/10.1038/s41593-019-0574-1)
- [17] T. P. Lillicrap, A. Santoro, L. Marris, C. J. Akerman, and G. Hinton, “Backpropagation and the brain,” *Nature Reviews Neuroscience*, vol. 21, no. 6, pp. 335–346, 2020, ISSN: 1471-003X, 1471-0048. DOI: [10.1038/s41583-020-0277-3](https://doi.org/10.1038/s41583-020-0277-3)
- [18] P. R. Roelfsema and A. Holtmaat, “Control of synaptic plasticity in deep cortical networks,” en, *Nature Reviews Neuroscience*, vol. 19, no. 3, pp. 166–180, 2018, Publisher: Nature Publishing Group, ISSN: 1471-0048. DOI: [10.1038/nrn.2018.6](https://doi.org/10.1038/nrn.2018.6)
- [19] E. R. KANDEL, “Dale’s Principle and the Functional Specificity of Neurons¹,” in *Psychopharmacology, a Review of Progress, 1957-1967: The Proceedings of the Sixth Annual Meeting of the American College of Neuropsychopharmacology, San Juan, Puerto Rico, December 12-15, 1967*, 1957, p. 385.
- [20] G. Buzsáki, “Hippocampal GABAergic Interneurons: A Physiological Perspective,” *Neurochemical Research*, vol. 26, no. 8, pp. 899–905, 2001, ISSN: 1573-6903. DOI: [10.1023/A:1012324231897](https://doi.org/10.1023/A:1012324231897)
- [21] H. Markram, M. Toledo-Rodriguez, Y. Wang, A. Gupta, G. Silberberg, and C. Wu, “Interneurons of the neocortical inhibitory system,” *Nature Reviews Neuroscience*, vol. 5, no. 10, pp. 793–807, 2004, ISSN: 1471-0048. DOI: [10.1038/nrn1519](https://doi.org/10.1038/nrn1519)

- [22] H. Hu, J. Gan, and P. Jonas, "Interneurons. Fast-spiking, parvalbumin⁺ GABAergic interneurons: From cellular design to microcircuit function," *Science (New York, N.Y.)*, vol. 345, no. 6196, p. 1255263, 2014, ISSN: 1095-9203. DOI: [10.1126/science.1255263](https://doi.org/10.1126/science.1255263)
- [23] F. Pouille and M. Scanziani, "Enforcement of Temporal Fidelity in Pyramidal Cells by Somatic Feed-Forward Inhibition," *Science*, vol. 293, no. 5532, pp. 1159–1163, 2001. DOI: [10.1126/science.1060342](https://doi.org/10.1126/science.1060342)
- [24] B. Amilhon et al., "Parvalbumin interneurons of hippocampus tune population activity at theta frequency," *Neuron*, vol. 86, no. 5, pp. 1277–1289, 2015, ISSN: 10974199. DOI: [10.1016/j.neuron.2015.05.027](https://doi.org/10.1016/j.neuron.2015.05.027)
- [25] P. Wulff et al., "Hippocampal theta rhythm and its coupling with gamma oscillations require fast inhibition onto parvalbumin-positive interneurons," *Proceedings of the National Academy of Sciences*, vol. 106, no. 9, pp. 3561–3566, 2009, ISSN: 0027-8424. DOI: [10.1073/pnas.0813176106](https://doi.org/10.1073/pnas.0813176106)
- [26] J. A. Cardin et al., "Driving fast-spiking cells induces gamma rhythm and controls sensory responses," *Nature*, vol. 459, no. 7247, pp. 663–667, 2009, ISSN: 1476-4687. DOI: [10.1038/nature08002](https://doi.org/10.1038/nature08002)
- [27] G. Buzsáki and X.-J. Wang, "Mechanisms of Gamma Oscillations," *Annual Review of Neuroscience*, vol. 35, no. Volume 35, 2012, pp. 203–225, 2012, ISSN: 0147-006X, 1545-4126. DOI: [10.1146/annurev-neuro-062111-150444](https://doi.org/10.1146/annurev-neuro-062111-150444)
- [28] N. Jeong et al., "Goal-specific hippocampal inhibition gates learning," *Nature*, pp. 1–9, 2025, ISSN: 1476-4687. DOI: [10.1038/s41586-025-08868-5](https://doi.org/10.1038/s41586-025-08868-5)
- [29] J. Urban-Ciecko and A. L. Barth, "Somatostatin-expressing neurons in cortical networks," *Nature Reviews Neuroscience*, vol. 17, no. 7, pp. 401–409, 2016, ISSN: 1471-003X, 1471-0048. DOI: [10.1038/nrn.2016.53](https://doi.org/10.1038/nrn.2016.53)
- [30] J. Waters, M. Larkum, B. Sakmann, and F. Helmchen, "Supralinear Ca²⁺ Influx into Dendritic Tufts of Layer 2/3 Neocortical Pyramidal Neurons In Vitro and In Vivo," *Journal of Neuroscience*, vol. 23, no. 24, pp. 8558–8567, 2003, ISSN: 0270-6474, 1529-2401. DOI: [10.1523/JNEUROSCI.23-24-08558.2003](https://doi.org/10.1523/JNEUROSCI.23-24-08558.2003)
- [31] M. E. Larkum, J. Waters, B. Sakmann, and F. Helmchen, "Dendritic Spikes in Apical Dendrites of Neocortical Layer 2/3 Pyramidal Neurons," *Journal of Neuroscience*, vol. 27, no. 34, pp. 8999–9008, 2007, ISSN: 0270-6474, 1529-2401. DOI: [10.1523/JNEUROSCI.1717-07.2007](https://doi.org/10.1523/JNEUROSCI.1717-07.2007)
- [32] M. E. Larkum, T. Nevian, M. Sandler, A. Polsky, and J. Schiller, "Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: A new unifying principle," *Science (New York, NY)*, vol. 325, no. 5941, pp. 756–760, 2009.
- [33] R. Tremblay, S. Lee, and B. Rudy, "GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits," *Neuron*, vol. 91, no. 2, pp. 260–292, 2016, ISSN: 0896-6273. DOI: [10.1016/j.neuron.2016.06.033](https://doi.org/10.1016/j.neuron.2016.06.033)
- [34] H.-J. Pi, B. Hangya, D. Kvitsiani, J. I. Sanders, Z. J. Huang, and A. Kepcs, "Cortical interneurons that specialize in disinhibitory control," *Nature*, vol. 503, no. 7477, pp. 521–524, 2013, ISSN: 1476-4687. DOI: [10.1038/nature12676](https://doi.org/10.1038/nature12676)
- [35] J. J. Letzkus et al., "A disinhibitory microcircuit for associative fear learning in the auditory cortex," *Nature*, vol. 480, no. 7377, pp. 331–335, 2011, ISSN: 1476-4687. DOI: [10.1038/nature10674](https://doi.org/10.1038/nature10674)
- [36] S. Zhang et al., "Long-range and local circuits for top-down modulation of visual cortex processing," *Science*, vol. 345, no. 6197, pp. 660–665, 2014. DOI: [10.1126/science.1254126](https://doi.org/10.1126/science.1254126)
- [37] J. Sacramento, R. Ponte Costa, Y. Bengio, and W. Senn, "Dendritic cortical microcircuits approximate the backpropagation algorithm," in *Advances in Neural Information Processing Systems*, vol. 31, Curran Associates, Inc., 2018.
- [38] N. Wagatsuma, S. Nobukawa, and T. Fukai, "A microcircuit model involving parvalbumin, somatostatin, and vasoactive intestinal polypeptide inhibitory interneurons for the modulation of neuronal oscillation during visual processing," *Cerebral Cortex*, vol. 33, no. 8, pp. 4459–4477, 2023, ISSN: 1047-3211. DOI: [10.1093/cercor/bhac355](https://doi.org/10.1093/cercor/bhac355)
- [39] J. Cornford, D. Kalajdzievski, M. Leite, A. Lamarquette, D. M. Kullmann, and B. Richards, *Learning to live with Dale's principle: ANNs with separate excitatory and inhibitory units*, 2021. DOI: [10.1101/2020.11.02.364968](https://doi.org/10.1101/2020.11.02.364968)
- [40] P. Li, J. Cornford, A. Ghosh, and B. Richards, "Learning better with Dale's Law: A Spectral Perspective," *Advances in Neural Information Processing Systems*, vol. 36, pp. 944–956, 2023.
- [41] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," 1986.
- [42] Y. Lecun, "A theoretical framework for back-propagation," *Proceedings of the 1988 Connectionist Models Summer School, CMU, Pittsburg, PA*, D. Touretzky, G. Hinton, and T. Sejnowski, Eds., pp. 21–28, 1988.

- [43] Y. Bengio, D.-H. Lee, J. Bornschein, T. Mesnard, and Z. Lin, *Towards Biologically Plausible Deep Learning*, 2016. DOI: [10.48550/arXiv.1502.04156](https://doi.org/10.48550/arXiv.1502.04156) [cs]. arXiv: [1502.04156](https://arxiv.org/abs/1502.04156) [cs].
- [44] B. A. Richards et al., “A deep learning framework for neuroscience,” *Nature neuroscience*, vol. 22, no. 11, pp. 1761–1770, 2019, ISSN: 1097-6256. DOI: [10.1038/s41593-019-0520-2](https://doi.org/10.1038/s41593-019-0520-2)
- [45] S. Grossberg, “Competitive learning: From interactive activation to adaptive resonance,” *Cognitive Science*, vol. 11, no. 1, pp. 23–63, 1987, ISSN: 0364-0213. DOI: [10.1016/S0364-0213\(87\)80025-3](https://doi.org/10.1016/S0364-0213(87)80025-3)
- [46] Q. Liao, J. Leibo, and T. Poggio, “How Important Is Weight Symmetry in Backpropagation?” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1, 2016, ISSN: 2374-3468. DOI: [10.1609/aaai.v30i1.10279](https://doi.org/10.1609/aaai.v30i1.10279)
- [47] A. Nøkland, “Direct feedback alignment provides learning in deep neural networks,” *Advances in neural information processing systems*, vol. 29, 2016.
- [48] Y. Bengio, *How Auto-Encoders Could Provide Credit Assignment in Deep Networks via Target Propagation*, 2014. DOI: [10.48550/arXiv.1407.7906](https://doi.org/10.48550/arXiv.1407.7906) arXiv: [1407.7906](https://arxiv.org/abs/1407.7906) [cs].
- [49] Y. Le Cun, “Learning Process in an Asymmetric Threshold Network,” in *Disordered Systems and Biological Organization*, E. Bienenstock, F. F. Soulié, and G. Weisbuch, Eds., Berlin, Heidelberg: Springer, 1986, pp. 233–240, ISBN: 978-3-642-82657-3. DOI: [10.1007/978-3-642-82657-3_24](https://doi.org/10.1007/978-3-642-82657-3_24)
- [50] D.-H. Lee, S. Zhang, A. Fischer, and Y. Bengio, *Difference Target Propagation*, 2015. DOI: [10.48550/arXiv.1412.7525](https://doi.org/10.48550/arXiv.1412.7525) arXiv: [1412.7525](https://arxiv.org/abs/1412.7525) [cs].
- [51] S. Bartunov, A. Santoro, B. A. Richards, L. Marrs, G. E. Hinton, and T. Lillicrap, “Assessing the Scalability of Biologically-Motivated Deep Learning Algorithms and Architectures,” *arXiv:1807.04587* [cs, stat], 2018. arXiv: [1807.04587](https://arxiv.org/abs/1807.04587) [cs, stat].
- [52] A. G. Ororbia and A. Mali, “Biologically Motivated Algorithms for Propagating Local Target Representations,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 4651–4658, 2019, ISSN: 2374-3468. DOI: [10.1609/aaai.v33i01.33014651](https://doi.org/10.1609/aaai.v33i01.33014651)
- [53] Y. Bengio, *Deriving Differential Target Propagation from Iterating Approximate Inverses*, 2020. DOI: [10.48550/arXiv.2007.15139](https://doi.org/10.48550/arXiv.2007.15139) arXiv: [2007.15139](https://arxiv.org/abs/2007.15139) [cs].
- [54] A. Meulemans, F. S. Carzaniga, J. A. K. Suykens, J. Sacramento, and B. F. Grewe, “A Theoretical Framework for Target Propagation,” *arXiv:2006.14331* [cs, stat], 2020. arXiv: [2006 . 14331](https://arxiv.org/abs/2006.14331) [cs, stat].
- [55] M. Ernoult et al., “Towards Scaling Difference Target Propagation by Learning Backprop Targets,” *arXiv:2201.13415* [cs], 2022. arXiv: [2201 . 13415](https://arxiv.org/abs/2201.13415) [cs].
- [56] T. Shibuya, N. Inoue, R. Kawakami, and I. Sato, *Fixed-Weight Difference Target Propagation*, 2022. DOI: [10.48550/arXiv.2212.10352](https://doi.org/10.48550/arXiv.2212.10352) arXiv: [2212 . 10352](https://arxiv.org/abs/2212.10352) [cs].
- [57] G. W. Lindsay, “Attention in Psychology, Neuroscience, and Machine Learning,” *Frontiers in Computational Neuroscience*, vol. 14, p. 29, 2020, ISSN: 1662-5188. DOI: [10.3389/fncom.2020.00029](https://doi.org/10.3389/fncom.2020.00029)
- [58] B. Choksi, M. Mozafari, C. B. O’May, B. Ador, A. Alamia, and R. VanRullen, “Brain-inspired predictive coding dynamics improve the robustness of deep neural networks,” in *NeurIPS 2020 Workshop SVRHM*, 2020.
- [59] R. P. N. Rao and D. H. Ballard, “Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects,” *Nature Neuroscience*, vol. 2, no. 1, pp. 79–87, 1999, ISSN: 1546-1726. DOI: [10.1038/4580](https://doi.org/10.1038/4580)
- [60] R. P. N. Rao, D. C. Gklezakos, and V. Sathish, “Active Predictive Coding: A Unifying Neural Model for Active Perception, Compositional Learning, and Hierarchical Planning,” *Neural Computation*, vol. 36, no. 1, pp. 1–32, 2023, ISSN: 0899-7667, 1530-888X. DOI: [10.1162/neco_a_01627](https://doi.org/10.1162/neco_a_01627)
- [61] N. Islah, G. Etter, M. Tugsbayar, T. Gurbuz, B. Richards, and E. Muller, *Learning to combine top-down context and feed-forward representations under ambiguity with apical and basal dendrites*, 2023. DOI: [10.48550/arXiv.2312.05484](https://doi.org/10.48550/arXiv.2312.05484) arXiv: [2312.05484](https://arxiv.org/abs/2312.05484) [q-bio].
- [62] L. B. Naumann, J. Keijser, and H. Sprekeler, “Invariant neural subspaces maintained by feedback modulation,” *eLife*, vol. 11, S. Ostojic and A. J. King, Eds., e76096, 2022, ISSN: 2050-084X. DOI: [10.7554/eLife.76096](https://doi.org/10.7554/eLife.76096)
- [63] D. O. Hebb, “The First Stage of Perception: Growth of the Assembly,” in *The Organization of Behavior*, Psychology Press, 1949, ISBN: 978-1-4106-1240-3.
- [64] T. V. P. Bliss and T. Lomo, “Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path,” *Journal of Physiology*, vol. 232, pp. 331–356, 1973.
- [65] H. Markram, “A network of tufted layer 5 pyramidal neurons,” *Cerebral Cortex*, vol. 7, no. 6, pp. 523–533, 1997, ISSN: 1047-3211. DOI: [10.1093/cercor/7.6.523](https://doi.org/10.1093/cercor/7.6.523)

- [66] G.-q. Bi and M.-m. Poo, "Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type," *Journal of Neuroscience*, vol. 18, no. 24, pp. 10 464–10 472, 1998, ISSN: 0270-6474, 1529-2401. DOI: [10.1523/JNEUROSCI.18-24-10464.1998](https://doi.org/10.1523/JNEUROSCI.18-24-10464.1998)
- [67] J. Lisman and N. Spruston, "Questions about STDP as a General Model of Synaptic Plasticity," *Frontiers in Synaptic Neuroscience*, vol. 2, 2010, ISSN: 1663-3563. DOI: [10.3389/fnsyn.2010.00140](https://doi.org/10.3389/fnsyn.2010.00140)
- [68] Y. Inglebert and D. Debanne, "Calcium and Spike Timing-Dependent Plasticity," *Frontiers in Cellular Neuroscience*, vol. 15, 2021, ISSN: 1662-5102.
- [69] G. Chindemi et al., "A calcium-based plasticity model for predicting long-term potentiation and depression in the neocortex," *Nature Communications*, vol. 13, no. 1, p. 3038, 2022, ISSN: 2041-1723. DOI: [10.1038/s41467-022-30214-w](https://doi.org/10.1038/s41467-022-30214-w)
- [70] K. C. Bittner et al., "Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons," *Nature Neuroscience*, vol. 18, no. July, pp. 1–13, 2015, ISSN: 1097-6256. DOI: [10.1038/nn.4062](https://doi.org/10.1038/nn.4062)
- [71] K. Bittner, A. Milstein, C. Grienberger, S. Romani, and J. Magee, "Behavioral time scale synaptic plasticity underlies CA1 place fields," *Science*, vol. 357, pp. 1033–1036, 2017. DOI: [10.1126/science.aan3846](https://doi.org/10.1126/science.aan3846)
- [72] A. D. Milstein et al., "Bidirectional synaptic plasticity rapidly modifies hippocampal representations," *eLife*, vol. 10, K. Toth, L. L. Colgin, J. Epsztein, and R. Naud, Eds., e73046, 2021, ISSN: 2050-084X. DOI: [10.7554/eLife.73046](https://doi.org/10.7554/eLife.73046)
- [73] C. C. Dorian, J. Taxidis, A. Arac, and P. Golshani, *Behavioral timescale synaptic plasticity in the hippocampus creates non-spatial representations during learning and is modulated by entorhinal inputs*, 2025. DOI: [10.1101/2024.08.27.609983](https://doi.org/10.1101/2024.08.27.609983)
- [74] G. J. Stuart and B. Sakmann, "Active propagation of somatic action potentials into neocortical pyramidal cell dendrites," *Nature*, vol. 367, no. 6458, pp. 69–72, 1994.
- [75] A. T. Gulleedge, S. B. Park, Y. Kawaguchi, and G. J. Stuart, "Heterogeneity of Phasic Cholinergic Signaling in Neocortical Neurons," *Journal of Neurophysiology*, vol. 97, no. 3, pp. 2215–2229, 2007, ISSN: 0022-3077. DOI: [10.1152/jn.00493.2006](https://doi.org/10.1152/jn.00493.2006)
- [76] M. J. Berridge, "Neuronal Calcium Signaling," *Neuron*, vol. 21, no. 1, pp. 13–26, 1998, ISSN: 0896-6273. DOI: [10.1016/S0896-6273\(00\)80510-3](https://doi.org/10.1016/S0896-6273(00)80510-3)
- [77] M. E. Larkum, K. M. M. Kaiser, and B. Sakmann, "Calcium electrogenesis in distal apical dendrites of layer 5 pyramidal cells at a critical frequency of back-propagating action potentials," *Proceedings of the National Academy of Sciences*, vol. 96, no. 25, pp. 14 600–14 604, 1999. DOI: [10.1073/pnas.96.25.14600](https://doi.org/10.1073/pnas.96.25.14600)
- [78] T. Nevian and B. Sakmann, "Spine Ca²⁺ Signaling in Spike-Timing-Dependent Plasticity," *Journal of Neuroscience*, vol. 26, no. 43, pp. 11 001–11 013, 2006, ISSN: 0270-6474, 1529-2401. DOI: [10.1523/JNEUROSCI.1749-06.2006](https://doi.org/10.1523/JNEUROSCI.1749-06.2006)
- [79] J. Lisman, "A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory," *Proceedings of the National Academy of Sciences*, vol. 86, no. 23, pp. 9574–9578, 1989. DOI: [10.1073/pnas.86.23.9574](https://doi.org/10.1073/pnas.86.23.9574)
- [80] R. C. Malenka and M. F. Bear, "LTP and LTD: An Embarrassment of Riches," *Neuron*, vol. 44, no. 1, pp. 5–21, 2004, ISSN: 0896-6273. DOI: [10.1016/j.neuron.2004.09.012](https://doi.org/10.1016/j.neuron.2004.09.012)
- [81] J. E. Lisman, "Bursts as a unit of neural information: Making unreliable synapses reliable," *Trends in Neurosciences*, vol. 20, no. 1, pp. 38–43, 1997, ISSN: 0166-2236, 1878-108X. DOI: [10.1016/S0166-2236\(96\)10070-9](https://doi.org/10.1016/S0166-2236(96)10070-9)
- [82] R. Naud et al., *Fast burst fraction transients convey information independent of the firing rate*, 2024. DOI: [10.1101/2022.10.07.511138](https://doi.org/10.1101/2022.10.07.511138)
- [83] R. Gütig, "Spiking neurons can discover predictive features by aggregate-label learning," *Science*, vol. 351, no. 6277, aab4113, 2016. DOI: [10.1126/science.aab4113](https://doi.org/10.1126/science.aab4113)
- [84] A. Payeur, J. Guerguiev, F. Zenke, B. A. Richards, and R. Naud, "Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits," *Nature Neuroscience*, vol. 24, no. 7, pp. 1010–1019, 2021, ISSN: 1546-1726. DOI: [10.1038/s41593-021-00857-x](https://doi.org/10.1038/s41593-021-00857-x)
- [85] M. McCloskey and N. J. Cohen, "Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem," in *Psychology of Learning and Motivation*, G. H. Bower, Ed., vol. 24, Academic Press, 1989, pp. 109–165. DOI: [10.1016/S0079-7421\(08\)60536-8](https://doi.org/10.1016/S0079-7421(08)60536-8)
- [86] R. Ratcliff, "Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions," *Psychological Review*, vol. 97, no. 2, pp. 285–308, 1990, ISSN: 1939-1471. DOI: [10.1037/0033-295X.97.2.285](https://doi.org/10.1037/0033-295X.97.2.285)
- [87] R. M. French, "Catastrophic forgetting in connectionist networks," *Trends in Cognitive Sciences*, vol. 3, no. 4, pp. 128–135, 1999, ISSN: 1364-6613, 1879-307X. DOI: [10.1016/S1364-6613\(99\)01294-2](https://doi.org/10.1016/S1364-6613(99)01294-2)

- [88] L. Wang, X. Zhang, H. Su, and J. Zhu, "A Comprehensive Survey of Continual Learning: Theory, Method and Application," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 8, pp. 5362–5383, 2024, ISSN: 1939-3539. DOI: [10.1109/TPAMI.2024.3367329](https://doi.org/10.1109/TPAMI.2024.3367329)
- [89] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, *Improving neural networks by preventing co-adaptation of feature detectors*, arXiv:1207.0580 [cs], 2012. DOI: [10.48550/arXiv.1207.0580](https://doi.org/10.48550/arXiv.1207.0580)
- [90] F. Sarfraz, E. Arani, and B. Zonooz, "Sparse Coding in a Dual Memory System for Lifelong Learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 8, pp. 9714–9722, 2023, ISSN: 2374-3468. DOI: [10.1609/aaai.v37i8.26161](https://doi.org/10.1609/aaai.v37i8.26161)
- [91] R. Yuste and T. Bonhoeffer, "Morphological Changes in Dendritic Spines Associated with Long-Term Synaptic Plasticity," *Annual Review of Neuroscience*, vol. 24, no. Volume 24, 2001, pp. 1071–1089, 2001, ISSN: 0147-006X, 1545-4126. DOI: [10.1146/annurev.neuro.24.1.1071](https://doi.org/10.1146/annurev.neuro.24.1.1071)
- [92] F. Zenke, B. Poole, and S. Ganguli, "Continual Learning Through Synaptic Intelligence," in *Proceedings of the 34th International Conference on Machine Learning*, PMLR, 2017, pp. 3987–3995.
- [93] J. Kirkpatrick et al., "Overcoming catastrophic forgetting in neural networks," en, *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. 3521–3526, 2017, arXiv:1612.00796 [cs], ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.1611835114](https://doi.org/10.1073/pnas.1611835114)
- [94] S. Dohare, J. F. Hernandez-Garcia, Q. Lan, P. Rahman, A. R. Mahmood, and R. S. Sutton, "Loss of plasticity in deep continual learning," en, *Nature*, vol. 632, no. 8026, pp. 768–774, 2024, Publisher: Nature Publishing Group, ISSN: 1476-4687. DOI: [10.1038/s41586-024-07711-7](https://doi.org/10.1038/s41586-024-07711-7)
- [95] W. E. Skaggs and B. L. McNaughton, "Replay of Neuronal Firing Sequences in Rat Hippocampus During Sleep Following Spatial Experience," *Science*, vol. 271, no. 5257, pp. 1870–1873, 1996. DOI: [10.1126/science.271.5257.1870](https://doi.org/10.1126/science.271.5257.1870)
- [96] A. K. Lee and M. A. Wilson, "Memory of Sequential Experience in the Hippocampus during Slow Wave Sleep," *Neuron*, vol. 36, no. 6, pp. 1183–1194, 2002, ISSN: 0896-6273. DOI: [10.1016/S0896-6273\(02\)01096-6](https://doi.org/10.1016/S0896-6273(02)01096-6)
- [97] Z. Nádasdy, H. Hirase, A. Czurkó, J. Csicsvari, and G. Buzsáki, "Replay and Time Compression of Recurring Spike Sequences in the Hippocampus," *Journal of Neuroscience*, vol. 19, no. 21, pp. 9497–9507, 1999, ISSN: 0270-6474, 1529-2401. DOI: [10.1523/JNEUROSCI.19-21-09497.1999](https://doi.org/10.1523/JNEUROSCI.19-21-09497.1999)
- [98] S. P. Jadhav, C. Kemere, P. W. German, and L. M. Frank, "Awake hippocampal sharp-wave ripples support spatial memory," *Science (New York, N.Y.)*, vol. 336, no. 6087, pp. 1454–1458, 2012, ISSN: 1095-9203. DOI: [10.1126/science.1217230](https://doi.org/10.1126/science.1217230)
- [99] G. Girardeau, K. Benchenane, S. I. Wiener, G. Buzsáki, and M. B. Zugardo, "Selective suppression of hippocampal ripples impairs spatial memory," *Nature Neuroscience*, vol. 12, no. 10, pp. 1222–1223, 2009, ISSN: 1546-1726. DOI: [10.1038/nn.2384](https://doi.org/10.1038/nn.2384)
- [100] I. Gridchyn, P. Schoenenberger, J. O'Neill, and J. Csicsvari, "Assembly-Specific Disruption of Hippocampal Replay Leads to Selective Memory Deficit," *Neuron*, vol. 106, no. 2, 291–300.e6, 2020, ISSN: 0896-6273. DOI: [10.1016/j.neuron.2020.01.021](https://doi.org/10.1016/j.neuron.2020.01.021)
- [101] R. Boyce, S. D. Glasgow, S. Williams, and A. Adamantidis, "Causal evidence for the role of REM sleep theta rhythm in contextual memory consolidation," *Science*, vol. 352, no. 6287, pp. 812–816, 2016, ISSN: 10959203. DOI: [10.1126/science.aad5252](https://doi.org/10.1126/science.aad5252)
- [102] M. A. A. van der Meer and D. Bendor, "Awake replay: Off the clock but on the job," *Trends in Neurosciences*, vol. 48, no. 4, pp. 257–267, 2025, ISSN: 0166-2236, 1878-108X. DOI: [10.1016/j.tins.2025.02.006](https://doi.org/10.1016/j.tins.2025.02.006)
- [103] A. Chaudhry et al., *On Tiny Episodic Memories in Continual Learning*, arXiv:1902.10486 [cs], 2019. DOI: [10.48550/arXiv.1902.10486](https://doi.org/10.48550/arXiv.1902.10486)
- [104] M. Riemer et al., *Learning to Learn without Forgetting by Maximizing Transfer and Minimizing Interference*, en, arXiv:1810.11910 [cs], 2019. DOI: [10.48550/arXiv.1810.11910](https://doi.org/10.48550/arXiv.1810.11910)
- [105] D. Lopez-Paz and M. A. Ranzato, "Gradient Episodic Memory for Continual Learning," in *Advances in Neural Information Processing Systems*, vol. 30, Curran Associates, Inc., 2017.
- [106] S.-A. Rebuffi, A. Kolesnikov, G. Sperl, and C. H. Lampert, "iCaRL: Incremental Classifier and Representation Learning," 2017, pp. 2001–2010.
- [107] R. Aljundi, M. Lin, B. Goujaud, and Y. Bengio, "Gradient based sample selection for online continual learning," in *Advances in Neural Information Processing Systems*, vol. 32, Curran Associates, Inc., 2019.
- [108] Z. Borsos, M. Mutny, and A. Krause, "coresets via Bilevel Optimization for Continual Learning and Streaming," in *Advances in Neural Information Processing Systems*, vol. 33, Curran Associates, Inc., 2020, pp. 14 879–14 890.

- [109] J. Yoon, D. Madaan, E. Yang, and S. J. Hwang, *Online Coreset Selection for Rehearsal-based Continual Learning*, arXiv:2106.01085 [cs], 2022. DOI: [10.48550/arXiv.2106.01085](https://doi.org/10.48550/arXiv.2106.01085)
- [110] D. Shim, Z. Mai, J. Jeong, S. Sanner, H. Kim, and J. Jang, “Online Class-Incremental Continual Learning with Adversarial Shapley Value,” en, *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 11, pp. 9630–9638, 2021, ISSN: 2374-3468. DOI: [10.1609/aaai.v35i11.17159](https://doi.org/10.1609/aaai.v35i11.17159)
- [111] R. Tiwari, K. Killamsetty, R. Iyer, and P. Shenoy, “GCR: Gradient Coreset Based Replay Buffer Selection for Continual Learning,” en, 2022, pp. 99–108.
- [112] A. Chaudhry, M. Ranzato, M. Rohrbach, and M. Elhoseiny, *Efficient Lifelong Learning with A-GEM*, arXiv:1812.00420 [cs], 2019. DOI: [10.48550/arXiv.1812.00420](https://doi.org/10.48550/arXiv.1812.00420)
- [113] H. Shin, J. K. Lee, J. Kim, and J. Kim, “Continual Learning with Deep Generative Replay,” in *Advances in Neural Information Processing Systems*, vol. 30, Curran Associates, Inc., 2017.
- [114] O. Ostapenko, M. Puscas, T. Klein, P. Jahnichen, and M. Nabi, “Learning to Remember: A Synaptic Plasticity Driven Framework for Continual Learning,” 2019, pp. 11 321–11 329.
- [115] L. Wang, B. Lei, Q. Li, H. Su, J. Zhu, and Y. Zhong, “Triple-Memory Networks: A Brain-Inspired Method for Continual Learning,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 5, pp. 1925–1934, 2022, ISSN: 2162-2388. DOI: [10.1109/TNNLS.2021.3111019](https://doi.org/10.1109/TNNLS.2021.3111019)
- [116] R. Kemker and C. Kanan, *FearNet: Brain-Inspired Model for Incremental Learning*, arXiv:1711.10563 [cs], 2018. DOI: [10.48550/arXiv.1711.10563](https://doi.org/10.48550/arXiv.1711.10563)
- [117] Y. Xiang, Y. Fu, P. Ji, and H. Huang, “Incremental Learning Using Conditional Adversarial Networks,” 2019, pp. 6619–6628.
- [118] Y. Cong, M. Zhao, J. Li, S. Wang, and L. Carin, “GAN Memory with No Forgetting,” in *Advances in Neural Information Processing Systems*, vol. 33, Curran Associates, Inc., 2020, pp. 16 481–16 494.
- [119] X. Liu et al., “Generative Feature Replay for Class-Incremental Learning,” 2020, pp. 226–227.
- [120] G. M. van de Ven, H. T. Siegelmann, and A. S. Tolias, “Brain-inspired replay for continual learning with artificial neural networks,” en, *Nature Communications*, vol. 11, no. 1, p. 4069, 2020, Publisher: Nature Publishing Group, ISSN: 2041-1723. DOI: [10.1038/s41467-020-17866-2](https://doi.org/10.1038/s41467-020-17866-2)
- [121] G. Etter, J. E. Carmichael, and S. Williams, “Linking temporal coordination of hippocampal activity to memory function,” *Frontiers in Cellular Neuroscience*, vol. 17, 2023, ISSN: 1662-5102. DOI: [10.3389/fncel.2023.1233849](https://doi.org/10.3389/fncel.2023.1233849)
- [122] R. Yuste, M. J. Gutnick, D. Saar, K. R. Delaney, and D. W. Tank, “Ca²⁺ accumulations in dendrites of neocortical pyramidal neurons: An apical band and evidence for two functional compartments,” *Neuron*, vol. 13, no. 1, pp. 23–43, 1994, ISSN: 0896-6273. DOI: [10.1016/0896-6273\(94\)90457-X](https://doi.org/10.1016/0896-6273(94)90457-X)
- [123] M. E. Larkum, J. J. Zhu, and B. Sakmann, “A new cellular mechanism for coupling inputs arriving at different cortical layers,” *Nature*, vol. 398, no. 6725, pp. 338–341, 1999, ISSN: 1476-4687. DOI: [10.1038/18686](https://doi.org/10.1038/18686)
- [124] R. Naud, Z. Friedenberger, and K. Toth, *Silences, Spikes and Bursts: Three-Part Knot of the Neural Code*, 2023. DOI: [10.48550/arXiv.2302.07206](https://doi.org/10.48550/arXiv.2302.07206) [q-bio].
- [125] M. E. J. Sheffield and D. A. Dombeck, “Calcium transient prevalence across the dendritic arbour predicts place field properties,” *Nature*, vol. 517, no. 7533, pp. 200–204, 2015, ISSN: 1476-4687. DOI: [10.1038/nature13871](https://doi.org/10.1038/nature13871)
- [126] M. E. Sheffield and D. A. Dombeck, “Dendritic mechanisms of hippocampal place field formation,” *Current Opinion in Neurobiology*, Neurobiology of Learning and Plasticity, vol. 54, pp. 1–11, 2019, ISSN: 0959-4388. DOI: [10.1016/j.conb.2018.07.004](https://doi.org/10.1016/j.conb.2018.07.004)
- [127] C. Mead, “Neuromorphic electronic systems,” *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1629–1636, 1990, ISSN: 1558-2256. DOI: [10.1109/5.58356](https://doi.org/10.1109/5.58356)
- [128] C. D. Schuman et al., *A Survey of Neuromorphic Computing and Neural Networks in Hardware*, 2017. DOI: [10.48550/arXiv.1705.06963](https://doi.org/10.48550/arXiv.1705.06963) arXiv: [1705.06963](https://arxiv.org/abs/1705.06963) [cs].
- [129] L. Chua, “Memristor-The missing circuit element,” *IEEE Transactions on Circuit Theory*, vol. 18, no. 5, pp. 507–519, 1971, ISSN: 2374-9555. DOI: [10.1109/TCT.1971.1083337](https://doi.org/10.1109/TCT.1971.1083337)
- [130] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, “The missing memristor found,” *Nature*, vol. 453, no. 7191, pp. 80–83, 2008, ISSN: 1476-4687. DOI: [10.1038/nature06932](https://doi.org/10.1038/nature06932)
- [131] D. Ielmini and H.-S. P. Wong, “In-memory computing with resistive switching devices,” *Nature Electronics*, vol. 1, no. 6, pp. 333–343, 2018, ISSN: 2520-1131. DOI: [10.1038/s41928-018-0092-2](https://doi.org/10.1038/s41928-018-0092-2)
- [132] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, “Nanoscale Memristor Device as Synapse in Neuromorphic Systems,” *Nano Letters*, vol. 10, no. 4, pp. 1297–1301, 2010, ISSN: 1530-6984, 1530-6992. DOI: [10.1021/nl904092h](https://doi.org/10.1021/nl904092h)

- [133] M. Prezioso, F. Merrikh-Bayat, B. D. Hoskins, G. C. Adam, K. K. Likharev, and D. B. Strukov, “Training and operation of an integrated neuromorphic network based on metal-oxide memristors,” *Nature*, vol. 521, no. 7550, pp. 61–64, 2015, ISSN: 1476-4687. DOI: [10.1038/nature14441](https://doi.org/10.1038/nature14441)
- [134] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, “Physics for neuromorphic computing,” *Nature Reviews Physics*, vol. 2, no. 9, pp. 499–510, 2020, ISSN: 2522-5820. DOI: [10.1038/s42254-020-0208-2](https://doi.org/10.1038/s42254-020-0208-2)
- [135] E. Strubell, A. Ganesh, and A. McCallum, “Energy and Policy Considerations for Modern Deep Learning Research,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 09, pp. 13 693–13 696, 2020, ISSN: 2374-3468. DOI: [10.1609/aaai.v34i09.7123](https://doi.org/10.1609/aaai.v34i09.7123)
- [136] D. Patterson et al., *Carbon Emissions and Large Neural Network Training*, 2021. DOI: [10.48550/arXiv.2104.10350](https://doi.org/10.48550/arXiv.2104.10350) arXiv: [2104.10350 \[cs\]](https://arxiv.org/abs/2104.10350).
- [137] M. Davies et al., “Loihi: A Neuromorphic Many-core Processor with On-Chip Learning,” *IEEE Micro*, vol. 38, no. 1, pp. 82–99, 2018, ISSN: 1937-4143. DOI: [10.1109/MM.2018.112130359](https://doi.org/10.1109/MM.2018.112130359)